Attorney's D[illegible] No.: 10559/229001 / P8794

APPLICATION

FOR

UNITED STATES LETTERS PATENT

TITLE: FORWARDING DATA IN A ROUTING ARCHITECTURE

APPLICANT: ANAND RANGARAIAN, SANJAY BAKSHI AND SATYENDRA YADAV

CERTIFICATE OF MAILING BY EXPRESS MAIL

Express Mail Label No. EL558600470US

I hereby certify that this correspondence is being deposited with the United States Postal Service as Express Mail Post Office to Addressee with sufficient postage on the date indicated below and is addressed to the Commissioner for Patents, Washington, D.C. 20231.

June 30, 2000
Date of Deposit

Signature

Derek W. Norwood
Typed or Printed Name of Person Signing Certificate

FORWARDING DATA IN A ROUTING ARCHITECTURE

BACKGROUND

The invention relates to data forwarding in a routing architecture.

A typical router has to process both packets that need to be forwarded and control packets to compute routing table. When a router forwards data, it looks up a destination IP (internet protocol) address in a routing table. The route lookup is a time consuming process and limits the performance of the single-box component. Other forwarding operations that limit the performance include validating a header checksum, decrementing a time-to-live, and recalculating and updating the header checksum. Checksum is a sum of a group of data items used for error checking. Time-to-live refers to time before a packet of data is deleted or discarded.

DESCRIPTION OF DRAWINGS

FIG. 1 illustrates a routing architecture according to an implementation of the invention.

FIG. 2 illustrates packet processing in a routing architecture according to an implementation of the invention.

DETAILED DESCRIPTION

An implementation of the invention may use a set of modularized routers and switches. In the modularized components, the network device control and management functions are separated from the packet forwarding functions.

A control component performs the control packet processing and a forwarding component performs the packet forwarding functions. Such modularized components provide flexible deployment in a network architecture.

The implementation of a routing architecture described below has multiple packet forwarding components. If each forwarding component were to perform a route lookup and other associated operations such as decrementing the time-to-live field, the use of resources would limit the performance of the routing architecture. Implementations of the invention reduce the number of route lookups and associated operations when a packet is forwarded through multiple packet forwarding components.

FIG. 1 shows a networking router architecture 100 with modularized components for flexible deployment of services. The router architecture 100 demonstrates a use of modularized components where the control and management functions of network devices are separated from the packet forwarding functions. A control component 200 for controlling and managing devices is connected to a first forwarding component 300. The first forwarding component 300, a second forwarding component 400, and a third forwarding component 500 are interconnected through Ethernet 800. As the name implies, the forwarding components 300, 400, and 500 function to forward information to an intended destination, such as a host moon 600. The architecture 100 shows the three forwarding components but there could be more or fewer, depending on the

topology of the destination in reference to a sender and a receiver of the information.

An implementation minimizes router lookups and other associated operations by using a mechanism that is called ARP (Address Resolution Protocol)-forwarding. There are three main aspects to ARP forwarding: routing table calculation, layer-2 forwarding, and proxy-ARP. These are explained as follows.

The control component 200 can calculate the routing table and alter the routing table for each packet forwarding component based on whether an egress-port is present in the forwarding component. An egress-port is defined as a port through which a packet leaves a router architecture for a destination. For example, for the architecture 100 as shown in FIG. 1, if the destination is the host moon 600, the egress-port would be an egress port 550 located in the third forwarding component 500. If the egress port is present in the particular forwarding component, no change is made to the table. If the egress-port is not present in the forwarding component, the control component 200 can change the table so that the ``egress port'' designated in the routing table is changed to a component's port that begins the best path to the egress-port. This port is called a proxy-egress-port. For example, a proxy-egress-port would be a port 350 located in the first forwarding component 300 because the egress-port is located not in the component 300 but in the third forwarding component 500. The best path would be calculated from the way

the packet forwarding components are interconnected called topology information maintained in the control component 200.

The forwarding components can forward requests and packets on layer-2 through internal ports of intermediate forwarding components, for instance, the intermediate component 400. Layer-2 refers to the second layer of a standard software layering architecture model (i.e. Open Systems Interconnection) where rules and protocols are compatible within the layer. An internal port in a forwarding component is a port that connects the forwarding component to another forwarding component within the same routing architecture. Internal ports 440 and 450 are located, for example, in the second forwarding component 400 in FIG. 1.

Proxy ARP is usually used for dial-up hosts. A remote access server serving a dial-up host responds to an ARP request meant for the dial-up host with the server's own MAC address. The MAC address is used to identify devices at the Media Access Control layer in a network architecture. In ARP-forwarding used in the implementation, an egress-component, defined as a component through which a packet leaves the networking router architecture, responds to an ARP-request coming in its internal port for a host or another device connected to one of egress-component's egress ports, by generating a proxy-ARP-reply with the egress-component's MAC address. An ingress component is a component through which a packet enters the networking router architecture. The egress-

component may communicate with the host on a directly connected line which may not support Ethernet.

ARP forwarding can function to minimize lookups as suggested in the following example: A packet enters a router architecture at an ingress-component, which can validate and modify at layer-3 a header of the packet and performs a route lookup to determine where the next destination and the egress-port are. The ingress-component broadcasts an ARP request for the hardware address of the destination or the next hop device of the egress-port. If the egress-port is present in the ingress component, the ARP-request is sent to the egress-port. The packet is sent to the destination when an ARP-reply comes back. But if the egress-port is not in the ingress component, the ARP-request is sent through the ingress-component's proxy-egress port. The intermediate forwarding components forward the request at layer-2 to enable the request to reach an egress-component with the egress-port. Since the request is a broadcast, it is forwarded from all ports except the one at which it entered.

Topology information of the forwarding components contained in a control component can be used to eliminate unnecessary broadcasts of the ARP request in multiple ports. After receiving the request, the egress-component replies with the MAC address of the egress-component. This ARP-reply is also forwarded at layer-2 by the intermediate components to the ingress-component. For the egress-component to know where to send back the ARP-reply, the ingress-component fills the

source IP field of the ARP request with the IP address of one of the non-proxy-egress ports if the proxy-egress-port does not have an IP address. Once the ARP-reply is received, the ingress-component encapsulates the original data packet with the egress-component's MAC address and transmits the packet through its proxy-egress-port. The intermediate components forward this packet also at layer-2 to the egress-component. The egress-component then routes the packet to the destination.

In the implementation described, the internal ports of the intermediate components perform a data link relay or, in other words, a bridging function. Furthermore, the intermediate components can act as transparent bridges, automatically and independently forwarding packets to bridge the ingress-component and the egress-component.

The above mechanism of ARP forwarding applied to the networking router architecture 100 is illustrated in FIG. 2. Assume that a packet originating from a host sun 700 is destined for the host moon 600. When a packet enters the router architecture 100 from the host sun 700 through an ingress-port 340 in phase 1, the first forwarding component 300 (the ingress-component) validates and modifies the layer-3 header of the packet and performs a route lookup by using its routing table 360 to determine the next destination and an egress-port. A port in the first forwarding component 300 is determined to be a proxy-egress port 350 and not an egress-port from which the packet can leave the architecture 100 for

the destination. In phase 2, the first forwarding component 300 broadcasts an ARP-request, which reaches the second forwarding component 400 through the proxy-egress-port 350. The request is a broadcast, and therefore, it is forwarded from all ports except where it came in. In phase 3, the second forwarding component 400 forwards the ARP request to the third forwarding component 500 (the egress-component) at layer-2. The second forwarding component 400 can act as a transparent bridge that can forward packets independently without communicating to either the host 600 or the host 700. In phase 4, the third forwarding component 500 responds to the request by sending its own MAC address from an egress port 550 to the second forwarding component 400. The ARP-reply containing the MAC address is forwarded at layer-2 to the first forwarding component 300 in phase 5. In phase 6, the first forwarding component 300 completes the layer-2 encapsulation of the packet with the MAC address and transmits it through the proxy-egress-port 350. The second forwarding component 400 forwards the encapsulated packet at layer 2 to the third forwarding component 500 in phase 7. The third forwarding component 500 then routes the packet to the destination, the host moon 600, in phase 8.

In the above disclosure, the packet is ``routed'' only at the first forwarding component and the third forwarding component and the route look up and other associated operations are performed only at the first and third forwarding components. The intermediate components act as

transparent bridges to forward the packet on layer 2. Thus, an application of ARP-forwarding to multiple forwarding components can reduce the route lookups to increase performance of the networking router architecture 100.

Implementations of the invention can be applied without modifying the existing protocols within the system. Any forwarding components that support routing and transparent bridging can be used to perform ARP forwarding to minimize route lookups.

The invention can be implemented, for example, in a computer program executable on a machine. The computer program can be stored on a storage medium, such as random access memory (RAM), readable by a general or special purpose programmable machine, for routing packets.

Other implementations are within the scope of the following claims.